Assessment: Applying the CRISP-DM Methodology

During this part of the session, you will select **one** of the following cases then document how you would apply the CRISP - Methodology using the following guide [Link] to answer the questions posed in each phase.

Cases

1. Mall Customer Segmentation Data

Problem

As the owner of Naivas Supermarket, you want to understand the customers who can quickly converge [Target Customers]. You will then report the insights from the data to the marketing team, who will later plan a strategy accordingly.

Context

You have some primary data about your customers through loyalty cards like Customer ID, age, gender, annual income, and spending score. You assign a spending score to the customer based on your defined parameters like customer behavior and purchasing data.

Additional Information

By the end of this case study, you will answer the following questions:

1. How can we achieve customer segmentation using a machine learning algorithm (KMeans Clustering)?
2. Who are the target customers with whom you can start a marketing strategy [easy to converse]?
3. How does the marketing strategy work in the real world?

Source: [Link]

2. Plant Seedlings

Problem

Can you differentiate a weed from a crop seedling? Use the given dataset to build a model to classify crop and weed seedlings. Deploy your model as a web app so farmers can use it.

Context

The Aarhus University Signal Processing group, in collaboration with the University of Southern Denmark, has recently released a dataset containing images of approximately 960 unique plants belonging to 12 species at several growth stages. This is what the group released as part of their release statement:

> Successful cultivation of maize depends mainly on the efficacy of weed control. Weed control during the first six to eight weeks after planting is crucial because weeds compete vigorously with the crop for nutrients and water during this period. Annual yield losses occur as a result of weed infestations in cultivated crops. Crop yield losses attributable to weeds vary with the type of weed, type of crop, and the environmental conditions involved. Generally, yield losses can vary from 10 to 100 %, depending on the level of weed control practiced. Rarely does one experience zero yield loss due to weeds... Yield losses occur as a result of weed interference with the crop's growth and development...This explains why effective weed control is imperative. To perform effective control, the first critical requirement is correct weed identification.

This dataset contains 5,539 images of crop and weed seedlings. The images are grouped into 12 classes, as shown in the above pictures. These classes represent common plant species in Danish agriculture. Each class contains RGB images that show plants at different growth stages. The images are in various sizes and are in png format.

Source: [Link], [Link]